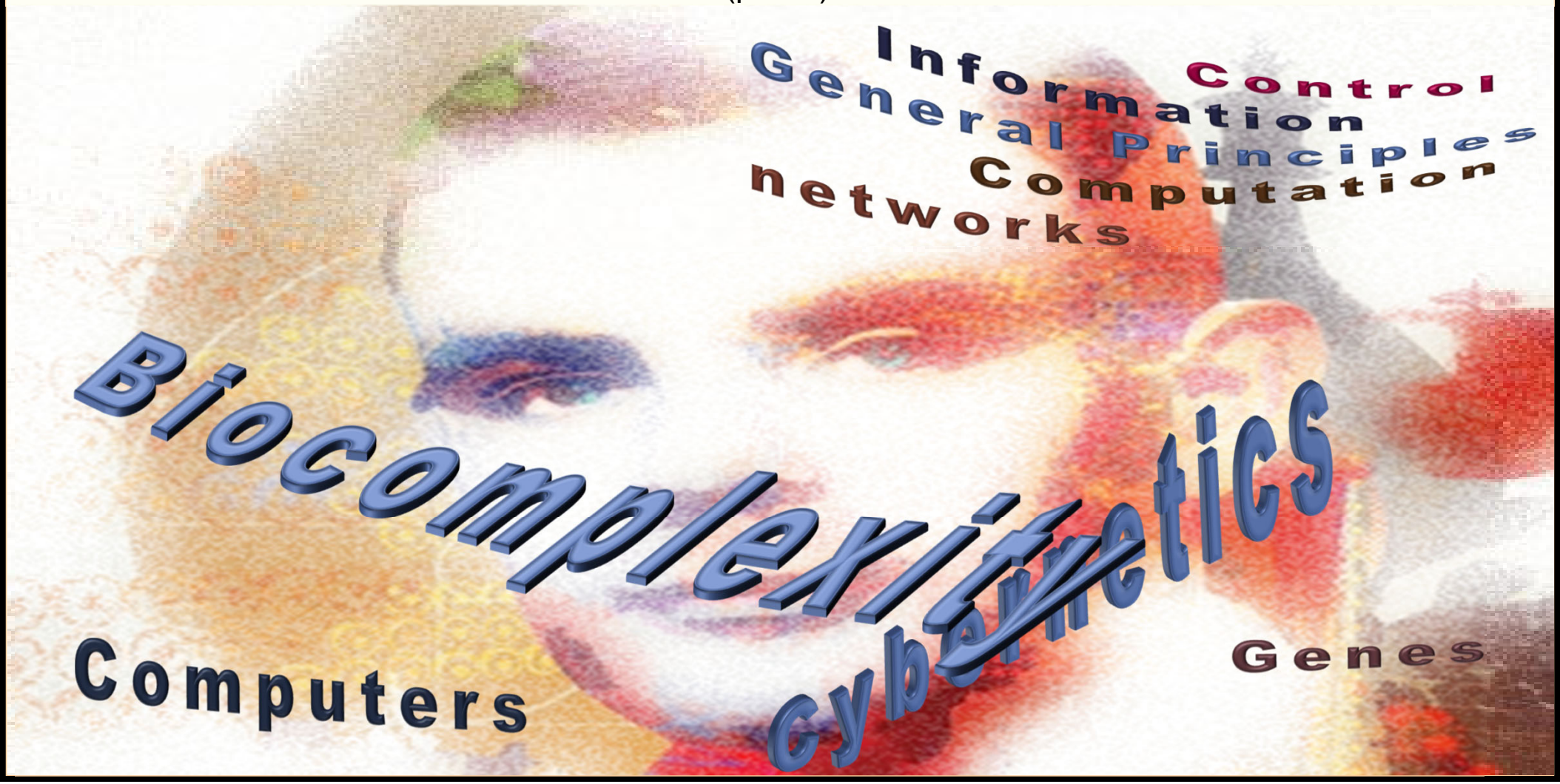
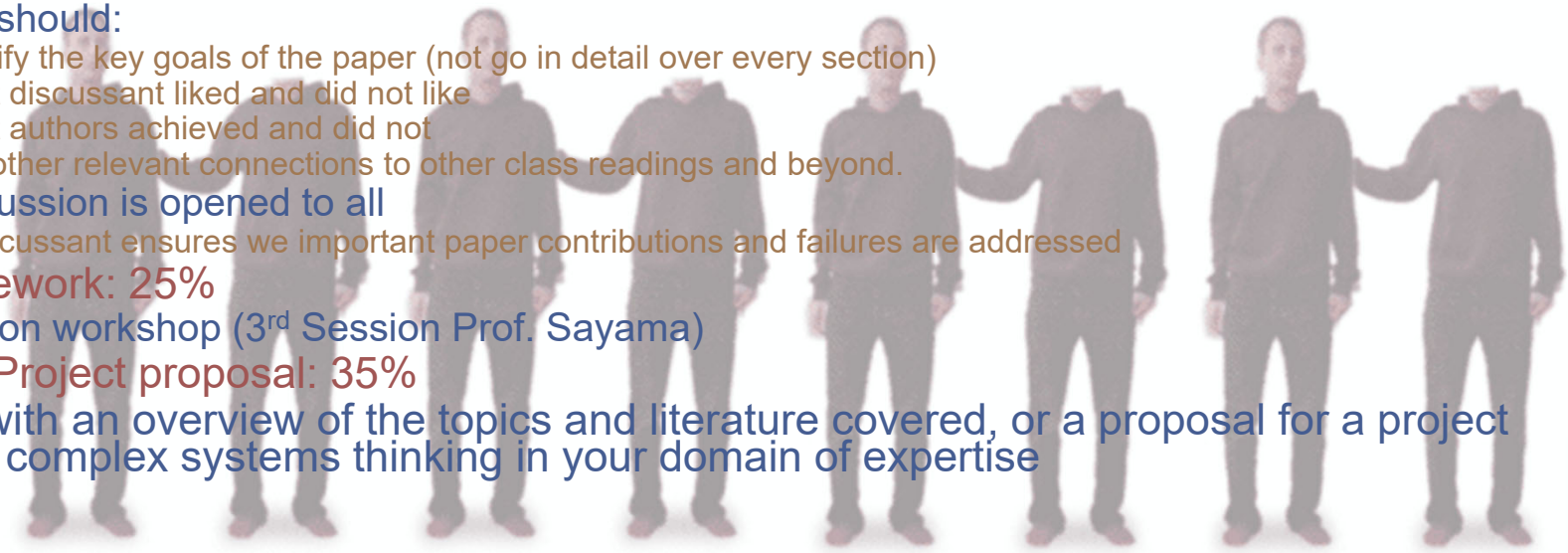
lecture 5: the information turn in the life sciences (part 2)



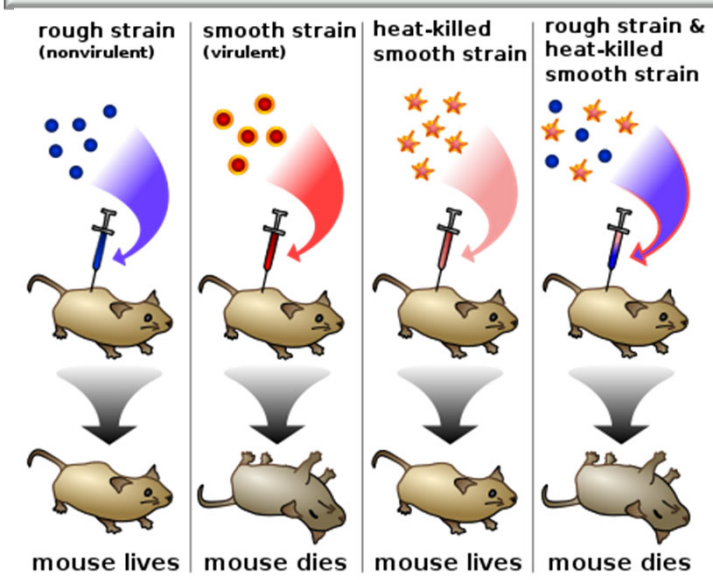
evaluation

- Participation and Discussion: 15%.
  - class discussion, everybody reads and discusses every paper
  - engagement in class
- Lead Discussions: 25%
  - Students are assigned to papers as lead discussants
    - all students are supposed to read and participate in discussion of every paper.
  - Lead discussant prepares short summary of assigned paper (10 minutes)
    - no formal presentations or PowerPoint unless figures are indispensable.
  - Summary should:
    - 1) Identify the key goals of the paper (not go in detail over every section)
    - 2) What discussant liked and did not like
    - 3) What authors achieved and did not
    - 4) Any other relevant connections to other class readings and beyond.
  - Class discussion is opened to all
    - lead discussant ensures we important paper contributions and failures are addressed
- Python Homework: 25%
  - From Python workshop (3rd Session Prof. Sayama)
- Term Paper/Project proposal: 35%
  - A paper with an overview of the topics and literature covered, or a proposal for a project that uses complex systems thinking in your domain of expertise



identifying the loci of genetic information

- Frederick Griffith's experiment
  - In 1928: Identified a "transforming principle"
- Avery's experiment
  - Oswald Avery, Colin MacLeod, and Maclyn McCarty
  - 1944: DNA as the loci of "transformation"
    - Chemically knocking off various cellular constituents until trying DNA
    - Considerable resistance in the community accepting this result until the early 1950's (Schrodinger, Delbruck, phage group)



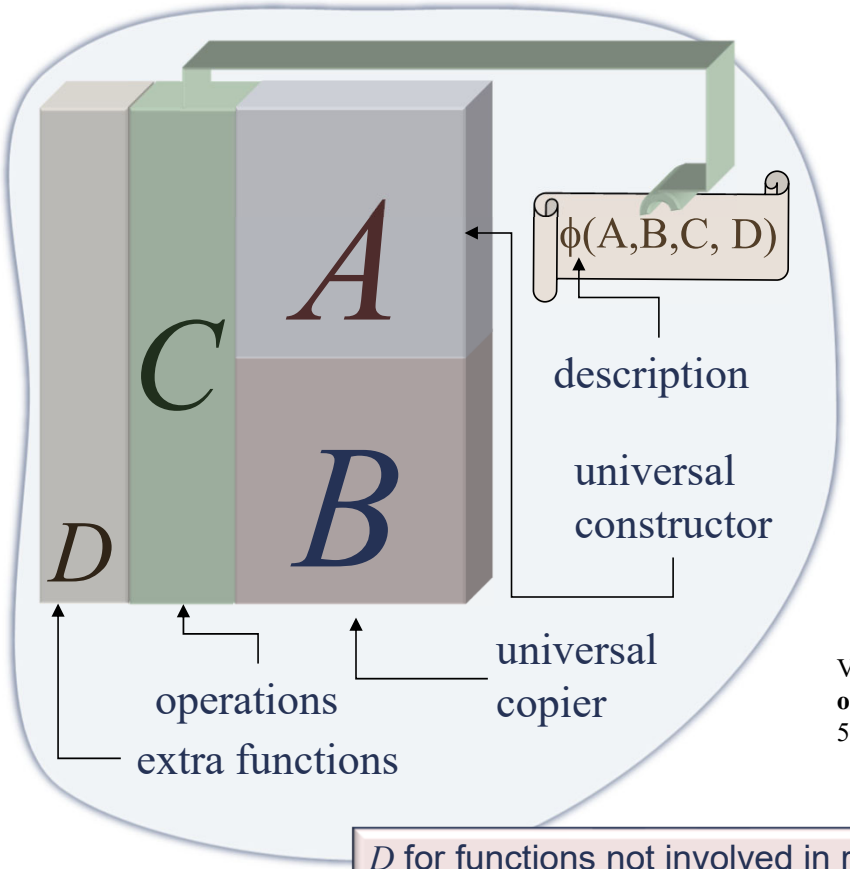
2 different strains of pneumococcus bacteria

initially not well accepted (No auto-catalysis with DNA)

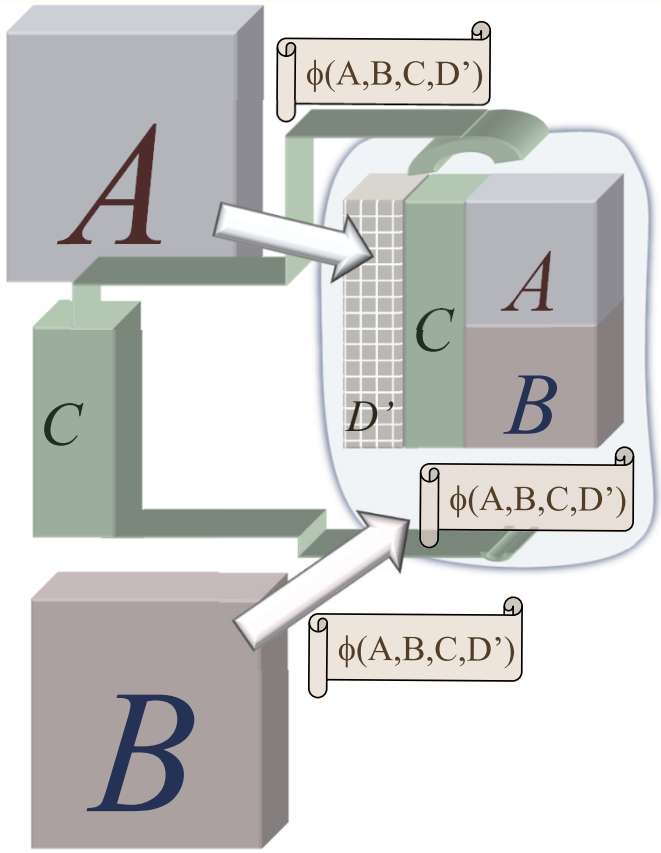
BINGHAMTON UNIVERSITY
STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501m

Von Neumann's generalization of Turing's tape

as a general principle (system) of evolution or **open-ended complexity**



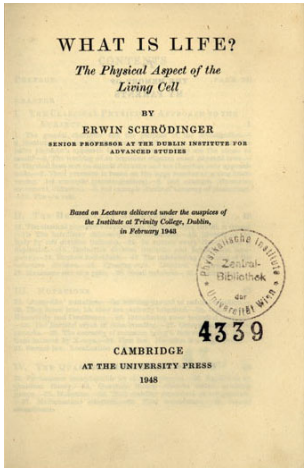
Von Neumann, J. [1949]. "**Theory and organization of complicated automata**." 5 lectures at University of Illinois



D for functions not involved in reproduction
Mutations in *D* can be propagated *vertically*
Leads to **open-ended evolution**

BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501m

Erwin Schrödinger(1943-1944)



- puzzled by the persistence of living structures
  - Call to understand how life stores and perpetuates order
  - "[…] **chromosomes**[…] contain in some kind of **code-script** the entire pattern of the individual's future development."
    - "complete (double) copy of the code-script."
- **aperiodic crystals as structures that can replicate themselves**
  - "We believe a gene—or perhaps the whole chromosome **fiber**—to be an aperiodic solid."
    - "structure without predictable repetition"
  - DNA is entirely regular
    - Instead of "aperiodicity" we have **encoded information**: separated description/construction

"Turing invented the stored-program computer, and von Neumann showed that the description is separate from the universal constructor. This is not trivial. Physicist Erwin Schrödinger confused the program and the constructor in his 1944 book *What is Life?,* in which he saw chromosomes as "*architect's plan and builder's craft in one*". This is wrong. The code script contains only a description of the executive function, not the function itself." (Sydney Brenner)

Brenner, Sydney. [2012]. "Life's code script." *Nature* **482** (7386): 461-461.

BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501m

Schrodinger vs. Von Neumann

self-replication vs. decoupled, encoded information

Von Neumann, J. [1949]. "**Theory and organization of complicated automata**." 5 lectures at University of Illinois

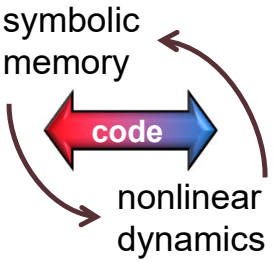
Brenner, Sydney. [2012]. "Life's code script." *Nature* **482** (7386): 461-461.

"Turing invented the stored-program computer, and von Neumann showed that the description is separate from the universal constructor. This is not trivial. Physicist Erwin Schrödinger confused the program and the constructor in his 1944 book *What is Life?,* in which he saw chromosomes as "*architect's plan and builder's craft in one*". This is wrong. The code script contains only a description of the executive function, not the function itself." (Sydney Brenner)

two roles of information
data/program (Turing)
passive/active (Von Neumann)
description/construction-function (Pattee)
genotype/phenotype (Biology)

semiotic closure (semiotic coupling)

fundamental principle of *organized complexity*
Leads to **open-ended evolution**
General principle that includes *Natural Selection*
Von Neumann described this scheme **before** structure of DNA molecule was identified in 1953 by Watson & Crick



Howard Pattee

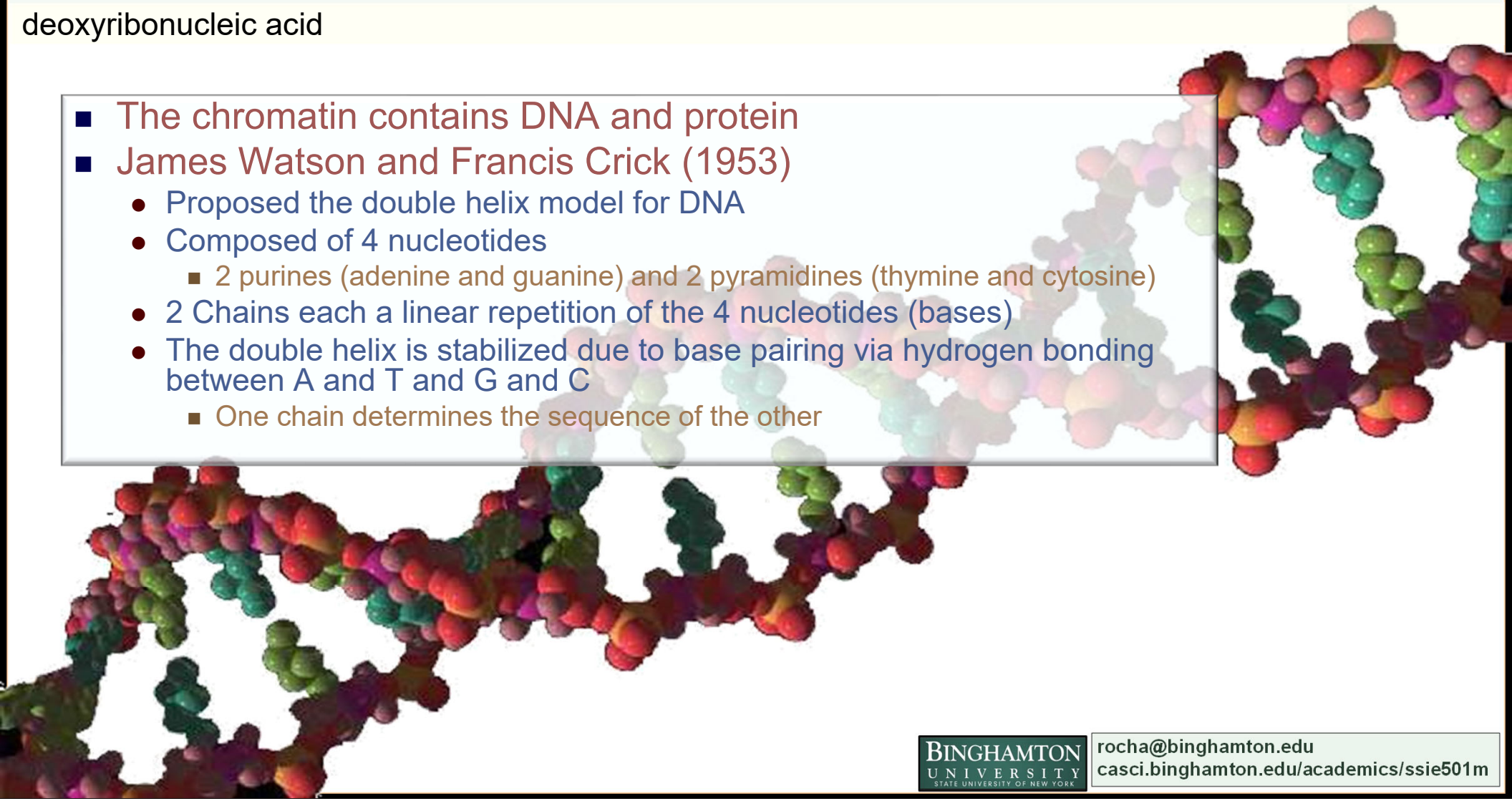
Pattee, HH [2001] *Biosystems* **60** (1):5-21

Rocha, L.M. & W. Hordijk [2005] *Artificial Life* **11**:189 - 214.
Rocha, L.M. [2001] *Biosystems* **60**: 95-121.
Rocha, L.M. [1996] *Systems Research* **13**: 371-384.

BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501m

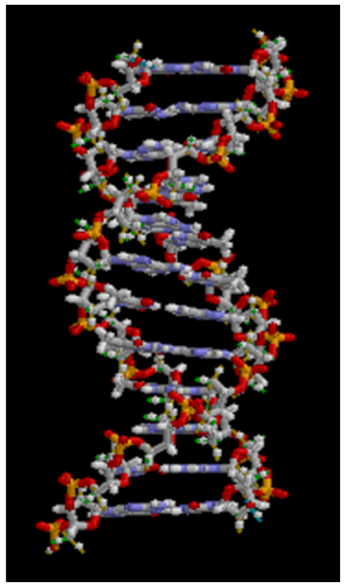
deoxyribonucleic acid

- The chromatin contains DNA and protein
- James Watson and Francis Crick (1953)
  - Proposed the double helix model for DNA
  - Composed of 4 nucleotides
    - 2 purines (adenine and guanine) and 2 pyramidines (thymine and cytosine)
  - 2 Chains each a linear repetition of the 4 nucleotides (bases)
  - The double helix is stabilized due to base pairing via hydrogen bonding between A and T and G and C
    - One chain determines the sequence of the other



BINGHAMTON UNIVERSITY
STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501m

a molecular language system: nucleotide “bases” (the genotype “tape”)



Purine (R) → Adenine (A), Guanine (G)

Nucleotides

Pyrimidine (Y) → Cytosine (C), Thymine (T), Uracil (U)

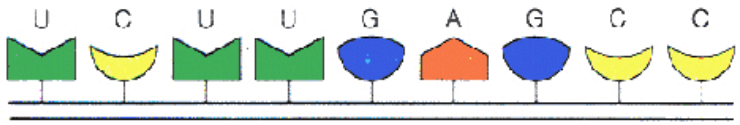
4 Letter Alphabet
DNA: A, G, C, T
RNA: A, G, C, U

Form sequences that can store information

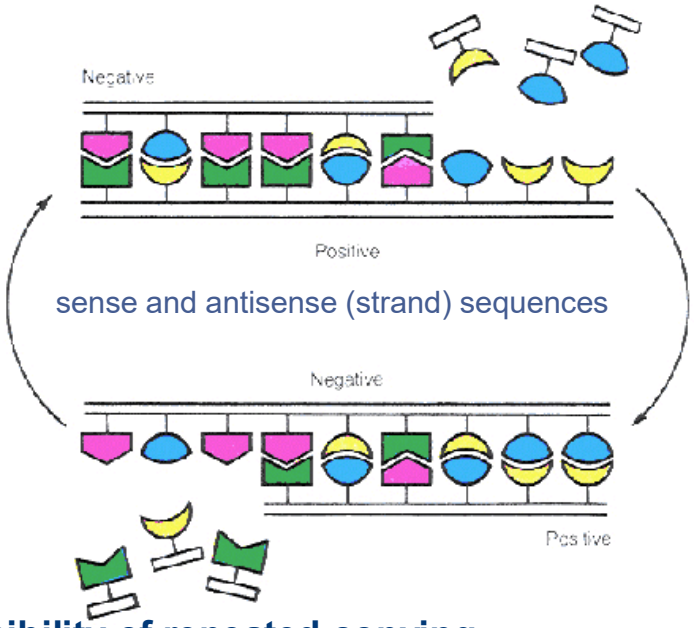
Linear molecules with a phosphate-sugar backbone (deoxyribose and ribose)

Complementary base pairing
(Hydrogen-bonding between purines and pyrimidines)

A-T (U) G-C



Figures from Eigen [1992] . *Steps Towards Life*.

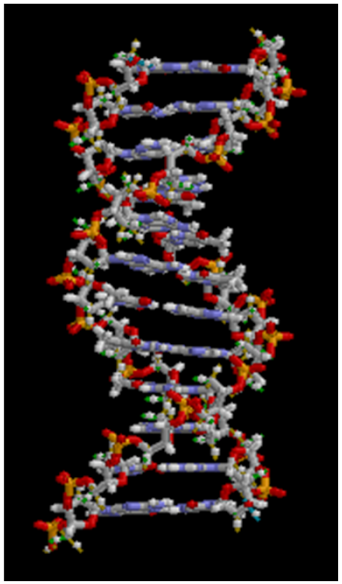


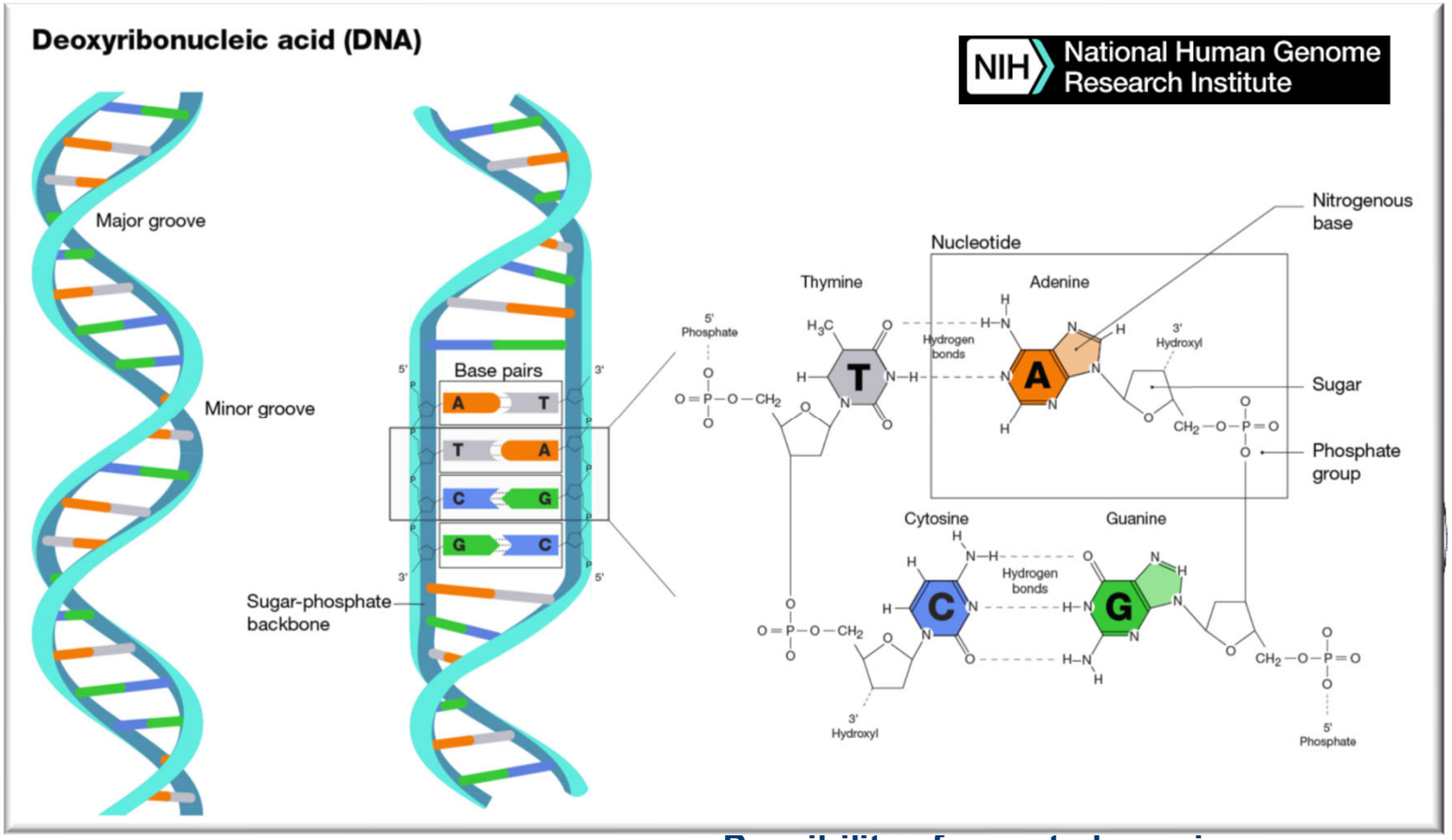
Requirements for structural information

Possibility of repeated copying

BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501m

a molecular language system: nucleotide “bases” (the genotype “tape”)

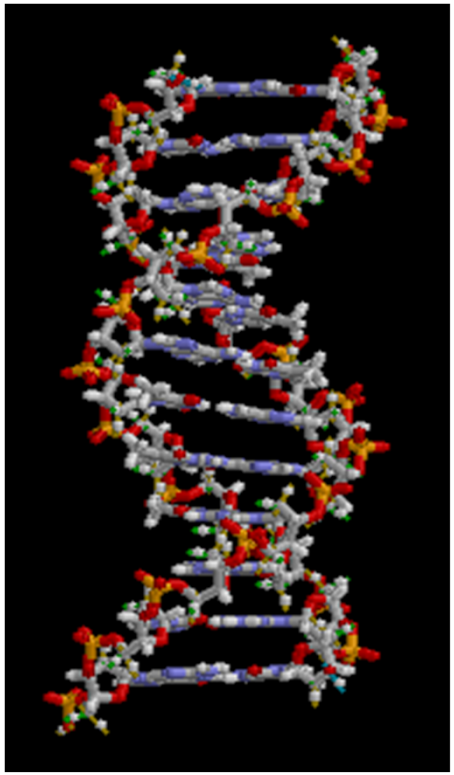


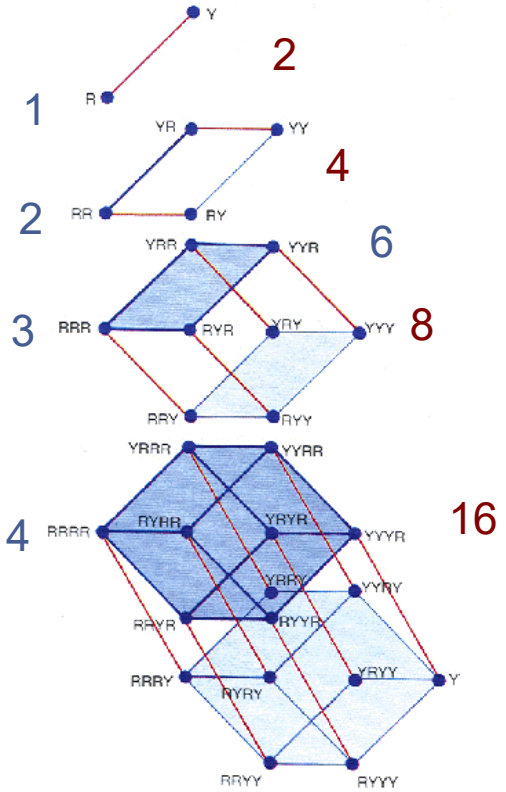


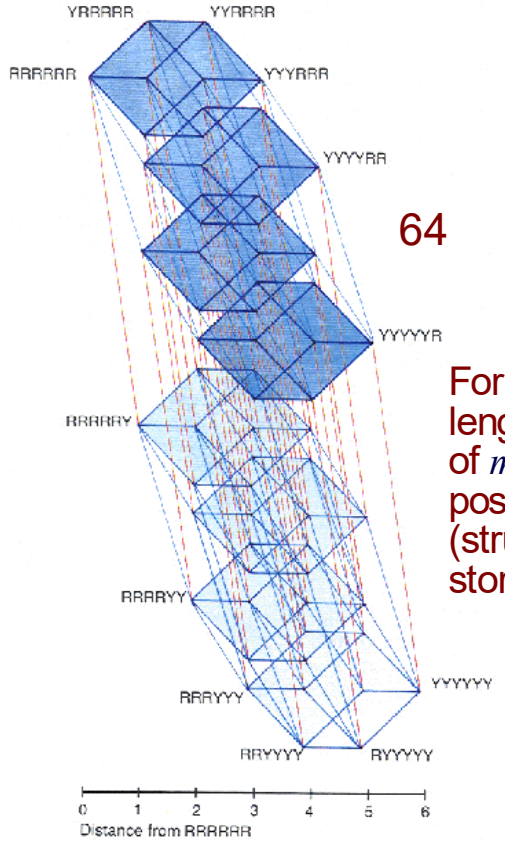
Possibility of repeated copying

BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501m

the genotype “tape” encodes an enormous amount of information





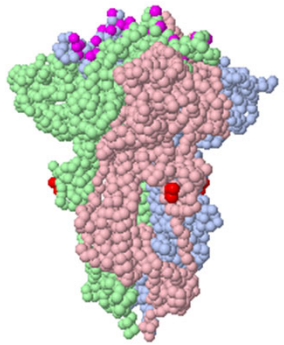


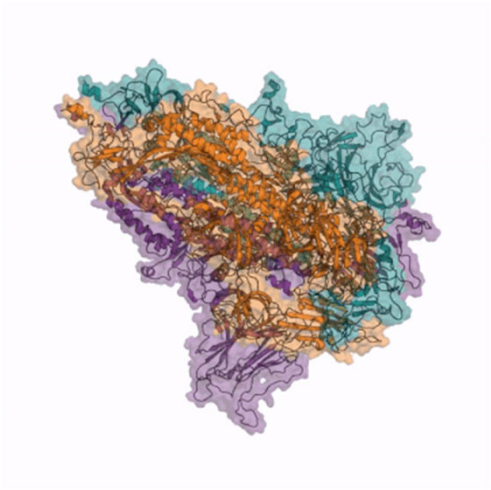
For a sequence of length n, composed of m-ary symbols, m^n possible values (structures) can be stored

Figures from Eigen [1992] . *Steps Towards Life*.

BINGHAMTON UNIVERSITY
STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501m

functional products that build up (self-organize) the phenotype





Polypeptide chains of aminoacids

Primary Structure

Folding

3-dimensional structure

Secondary and tertiary bonds

- In proteins, it is the 3-dimensional structure that dictates function
  - The specificity of enzymes to recognize and react on substrates
- The functioning of the cell is mostly performed by proteins
  - Though there are also ribozymes

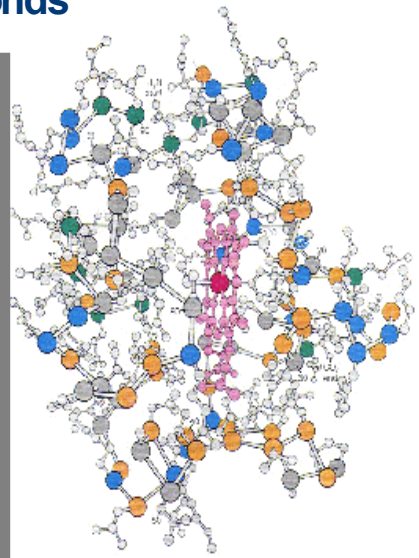


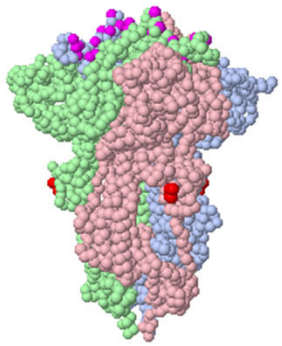
Table 1.4. Amino acid codes

Ala	A	Alanine
Arg	R	Arginine
Asn	N	Asparagine
Asp	D	Aspartic acid
Cys	C	Cysteine
Gln	Q	Glutamine
Glu	E	Glutamic acid
Gly	G	Glycine
His	H	Histidine
Ile	I	Isoleucine
Leu	L	Leucine
Lys	K	Lysine
Met	M	Methionine
Phe	F	Phenylalanine
Pro	P	Proline
Ser	S	Serine
Thr	T	Threonine
Trp	W	Tryptophan
Tyr	Y	Tyrosine
Val	V	Valine
Asx	B	Asn or Asp
Glx	Z	Gln or Glu
Sec	U	Selenocysteine
Unk	X	Unknown

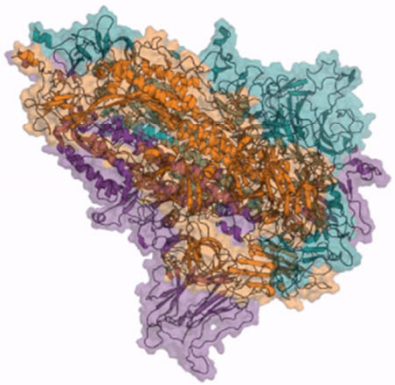
Figures from Eigen [1992] . *Steps Towards Life*.

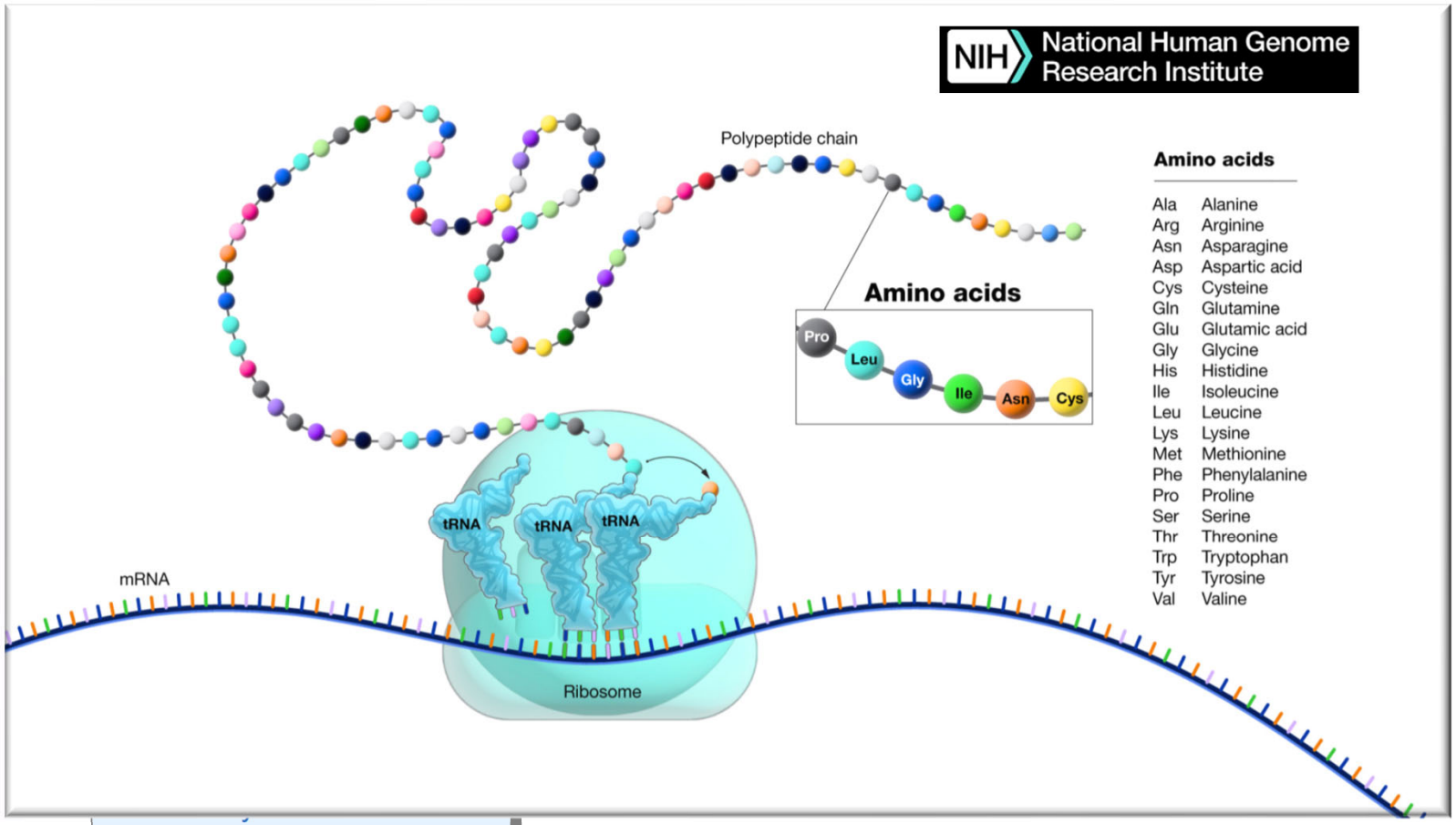
BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501m

functional products that build up (self-organize) the phenotype



Jmol

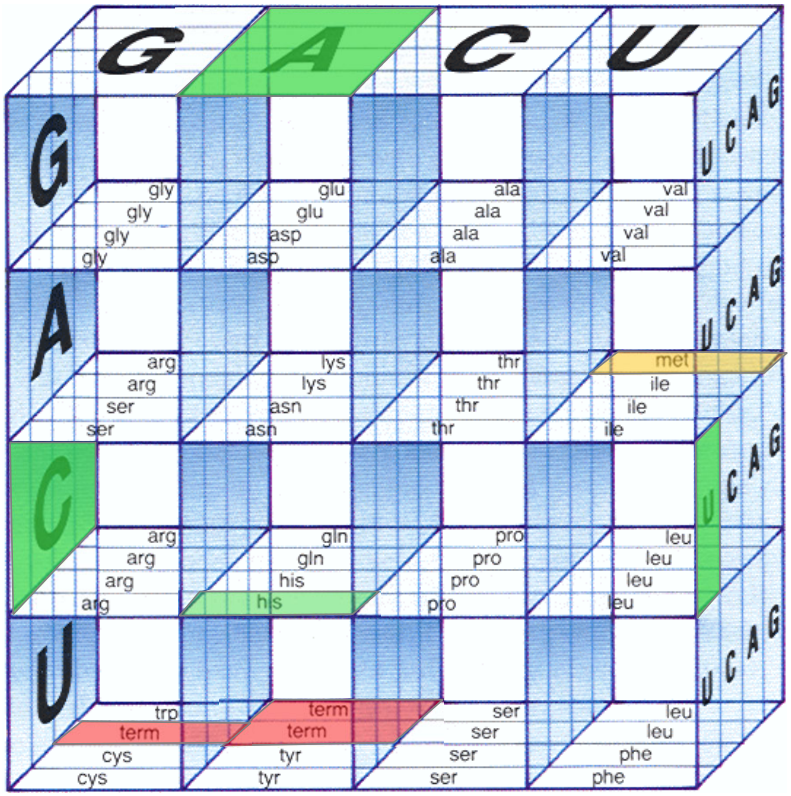




BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501m

between genotype and phenotype

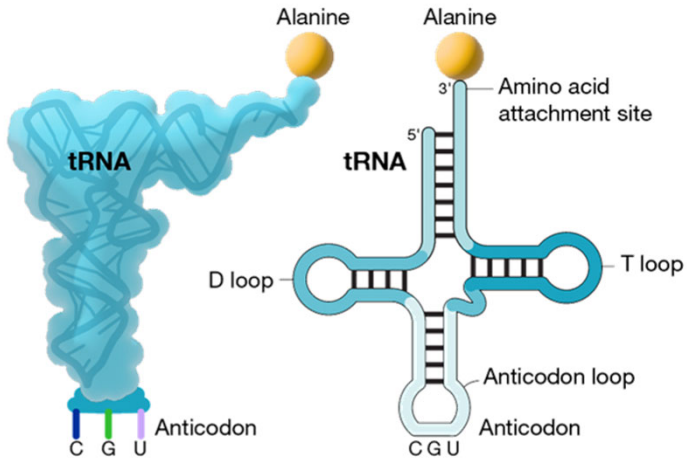
Triplets of 3 Nucleotides can define 64 possible codons, but only 20 amino acids are used (redundancy)



Figures from Eigen [1992] . *Steps Towards Life*.

- The genetic code maps information stored in the genome into functional proteins
  - Triplet combinations of nucleotides into amino acids

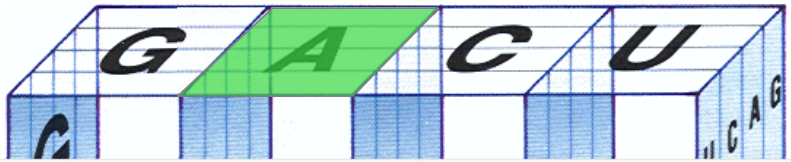
Common ways of depicting transfer RNA (tRNA)



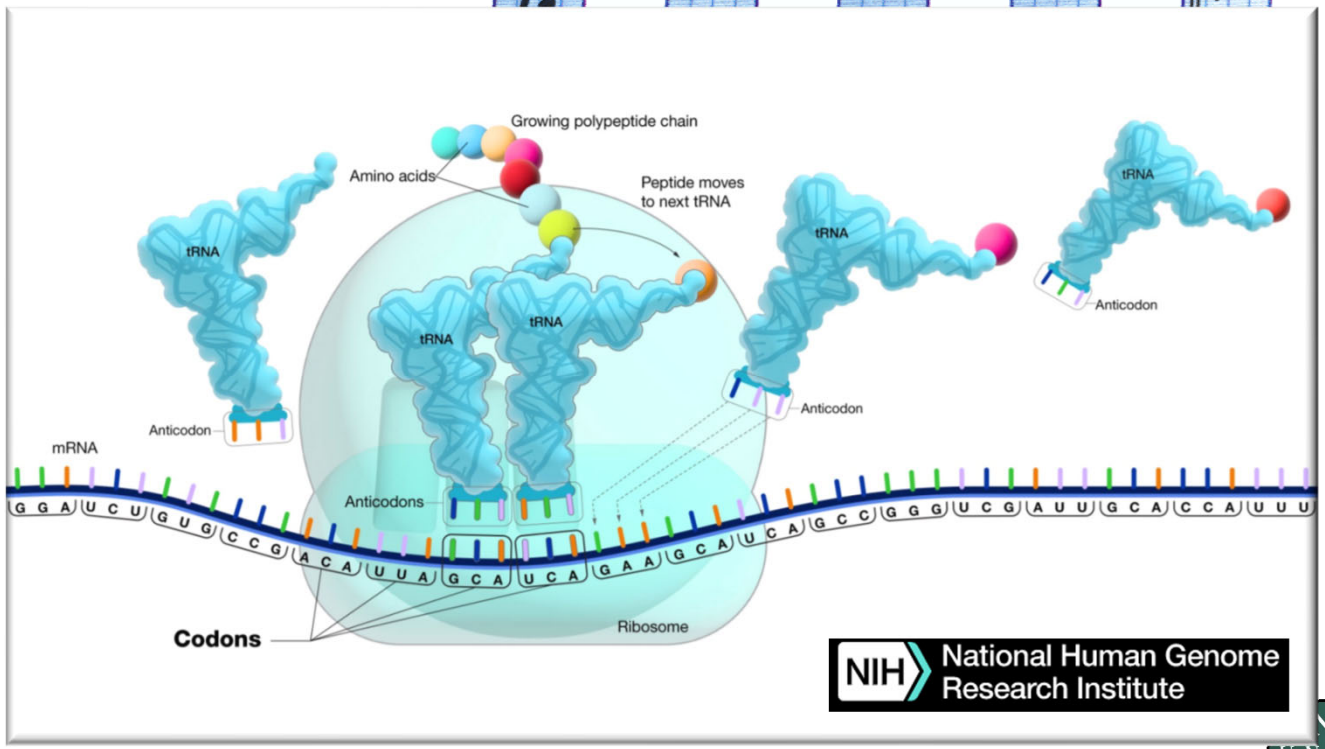
BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501m

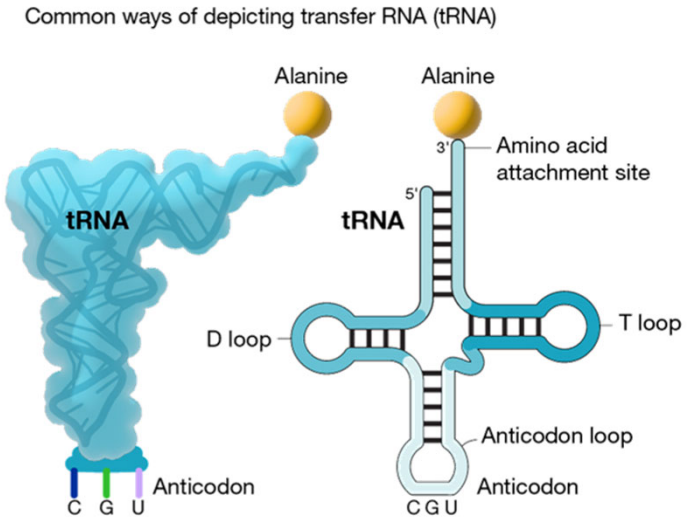
between genotype and phenotype

Triplets of 3 Nucleotides can define 64 possible codons, but only 20 amino acids are used (redundancy)



- The genetic code maps information stored in the genome into functional proteins
  - Triplet combinations of nucleotides into amino acids

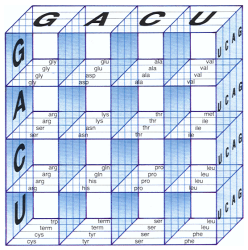




BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501m

semiotic closure: genetic information at work

genotype/phenotype



Genotype

DNA

Memory

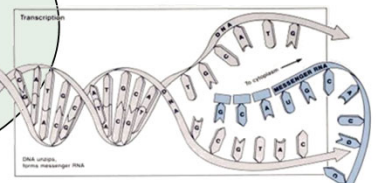
C

C

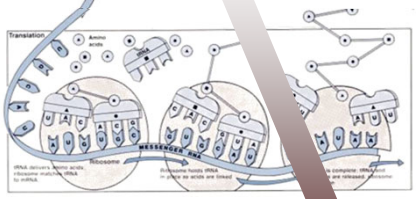
syntax

transcription

RNA

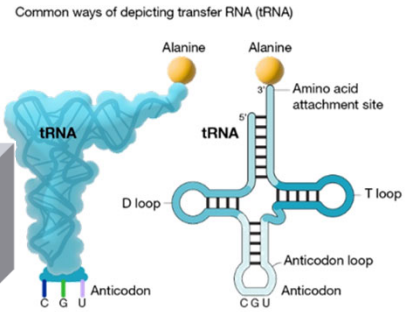


translation



code

A



semantics

amino acid chains

Development, regulation

pragmatics

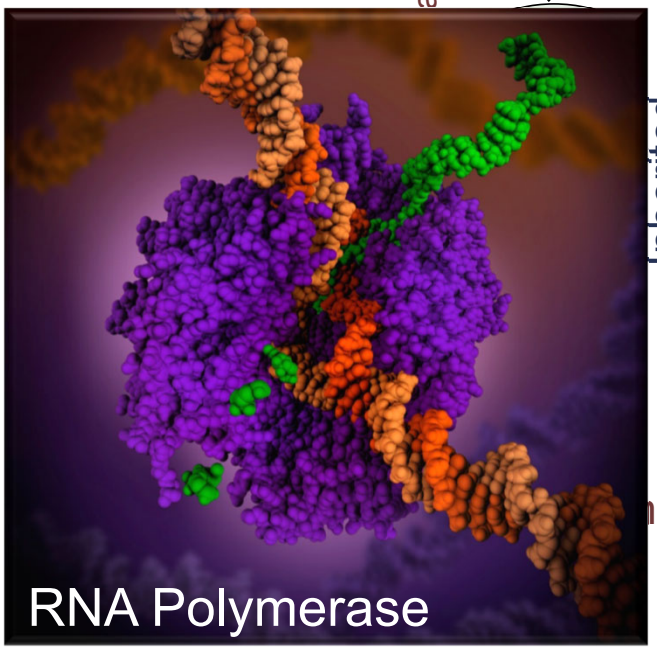
D

environmental ramifications

selection

Inherited variation

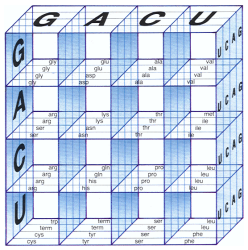
cell line



BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501m

semiotic closure: genetic information at work

genotype/phenotype



Genotype

DNA

Memory

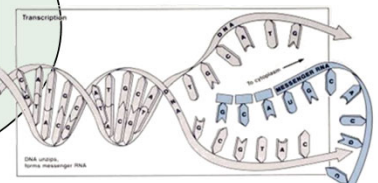
C

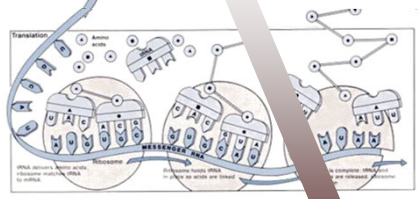
syntax

transcription

RNA

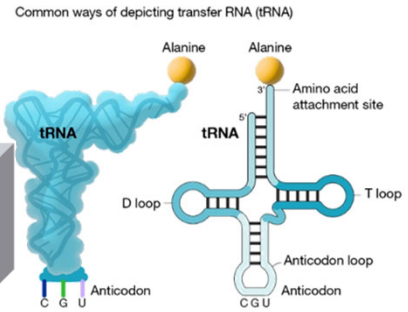
translation





code

A



semantics

C

replication

B

Inherited variation

selection

amino acid chains

Development regulation

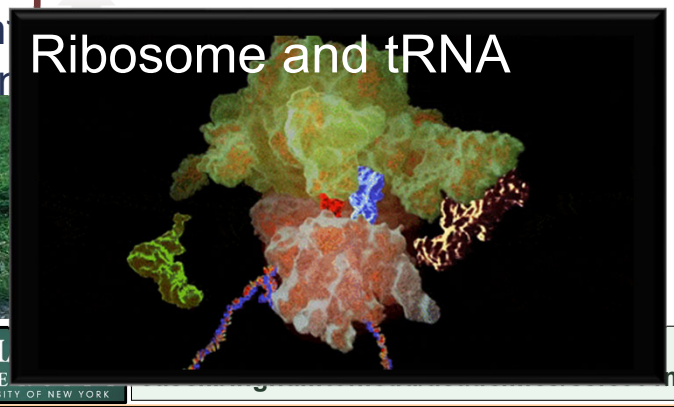
pragmatics

Germ cell line

D

environmental ramifications

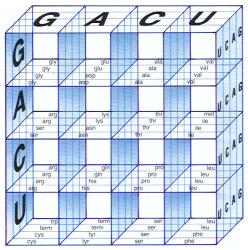
Ribosome and tRNA

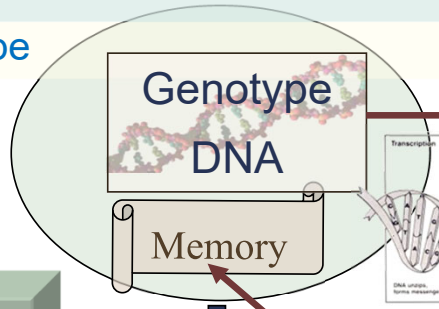


BINGH
UNIVE
STATE UNIVERSITY OF NEW YORK

semiotic closure: genetic information at work

genotype/phenotype

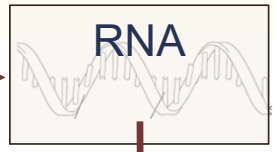


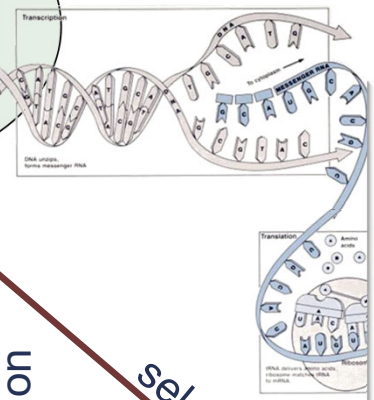


C

syntax

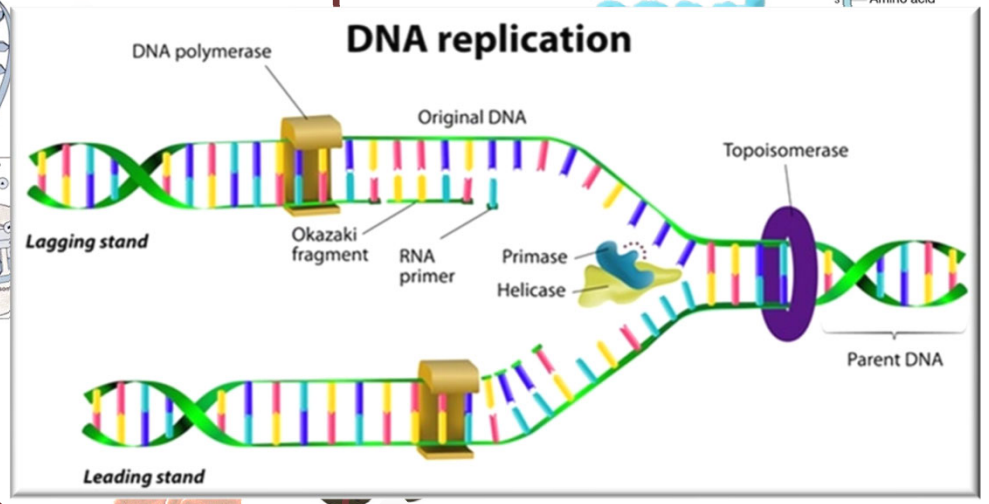
transcription

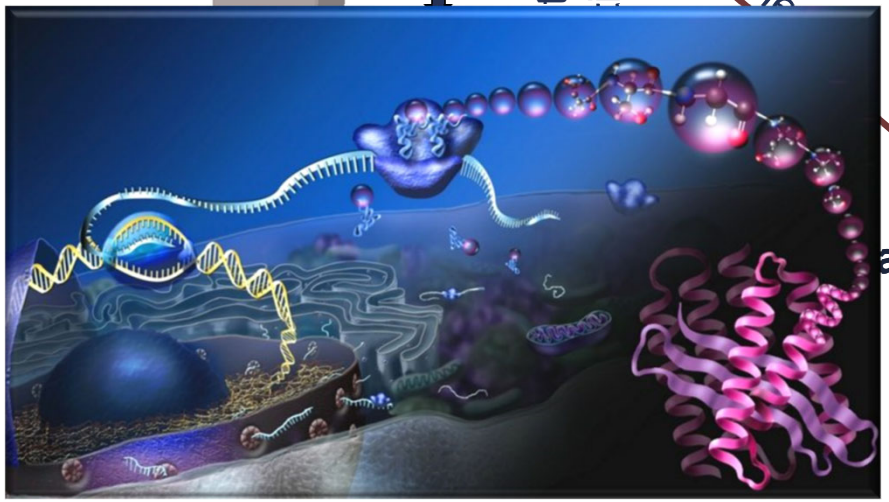




C

replication





atics

Development, regulation

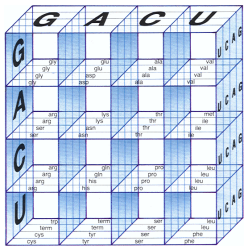
D

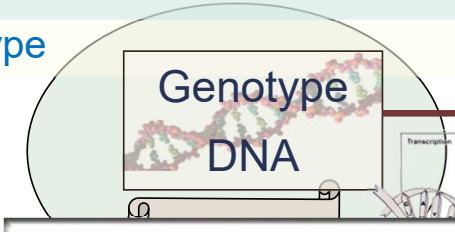
environmental ramifications

BINGHAMTON UNIVERSITY
STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501m

semiotic closure: genetic information at work

genotype/phenotype

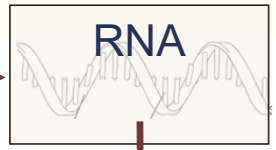


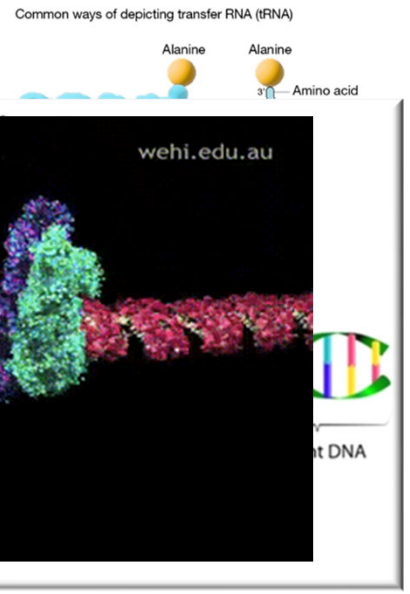


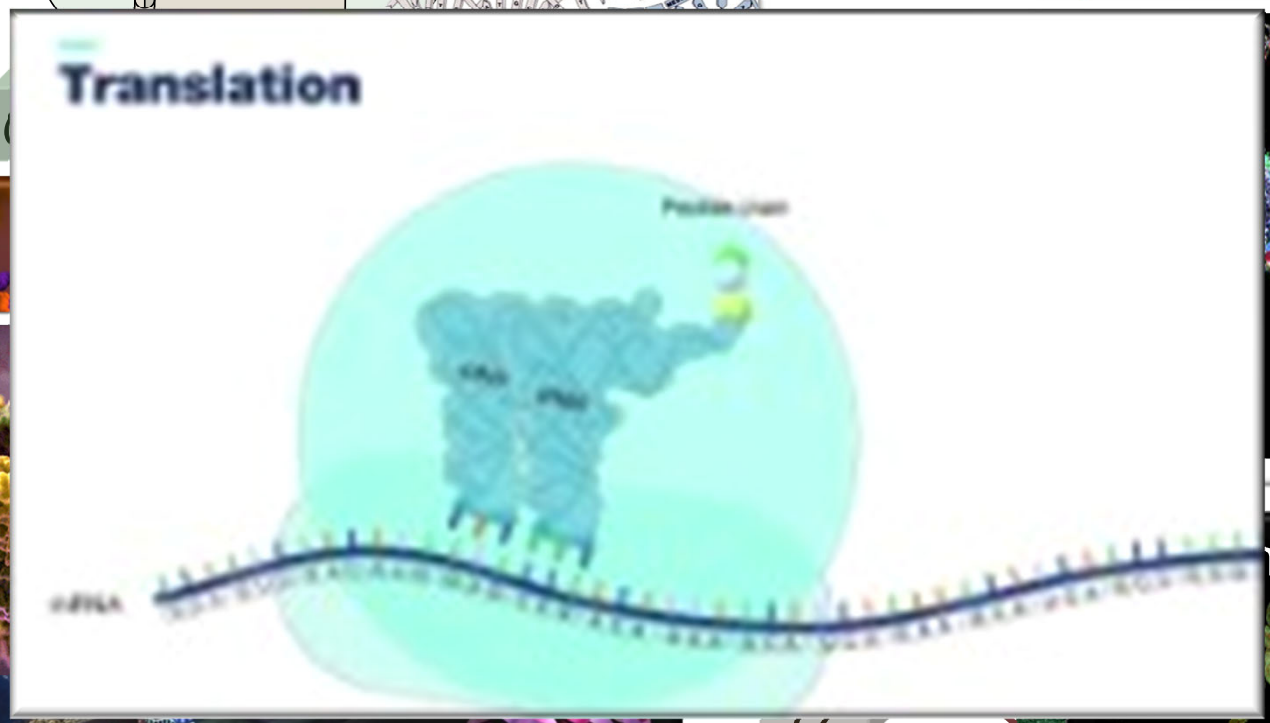
C

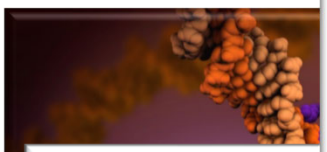
syntax

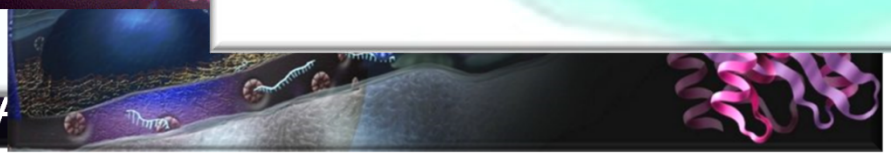
transcription





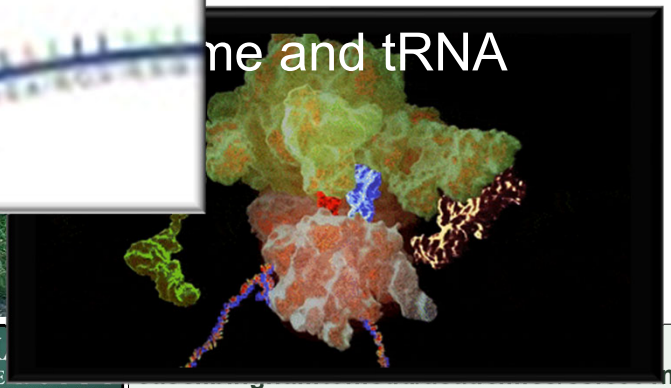






RNA

environmental
ramifications



BINGH
UNIVE
STATE UNIVERSITY OF NEW YORK

importance of the “external tape”

in biology

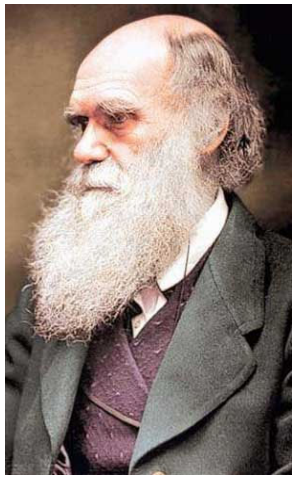
- The “information turn”
  - Unlike Schrödinger, Turing and Von Neumann had no direct effect on molecular biology
  - But the “external tape” separated from the constructor (semiotic closure) has become an unavoidable **principle of organization of biocomplexity**
  - A new synthesis?
    - In 1971 Brenner: “in the next twenty-five years we are going to have to teach biologists another language still, […] where a science like physics works in terms of laws, or a science like molecular biology, to now, is stated in terms of mechanisms, maybe now what one has to begin to think of is algorithms. Recipes. Procedures.”

“The concept of the gene as a symbolic representation of the organism — a ***code script*** — is a fundamental feature of the living world and must form the kernel of biological theory. […] at the core of everything are the tapes containing the descriptions to build these special Turing machines.” (Sydney Brenner)

Brenner, Sydney. [2012]. “Life’s code script.” *Nature* **482** (7386): 461-461.

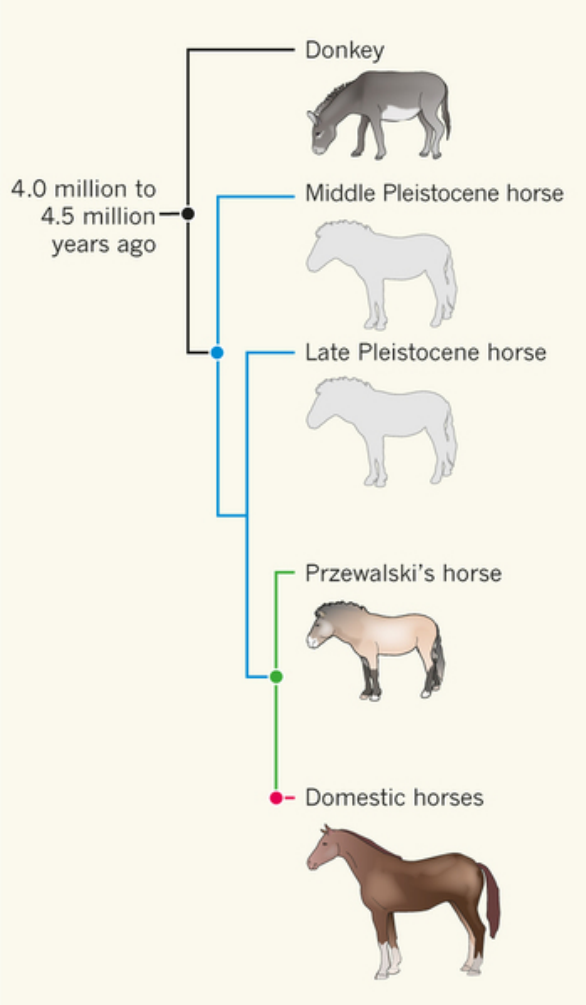
BINGHAMTON UNIVERSITY
STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501m

fundamental principle of organisms as *cybernetic mechanisms*



BINGHAMTON UNIVERSITY
STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501m

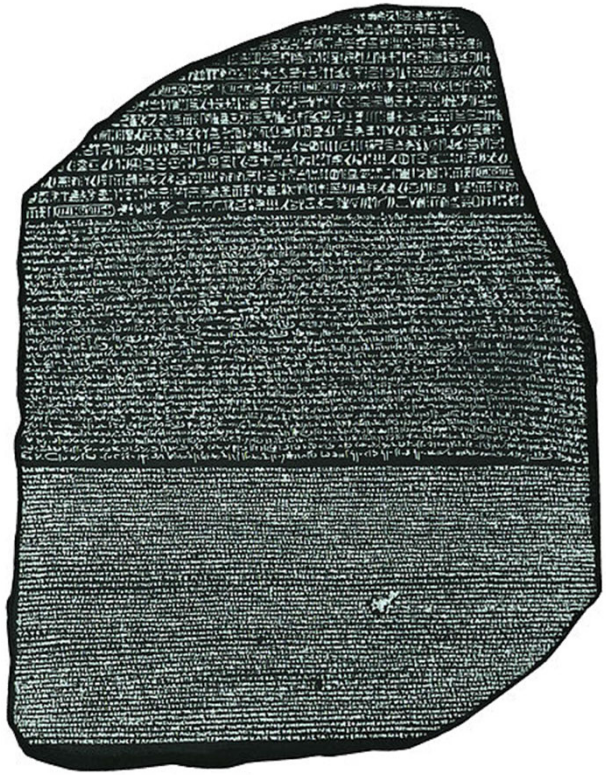
decoupled information



Millar & Lambert [2013]. "Ancient DNA: Towards a million-year-old genome." *Nature*. **doi:10.1038/nature12263**

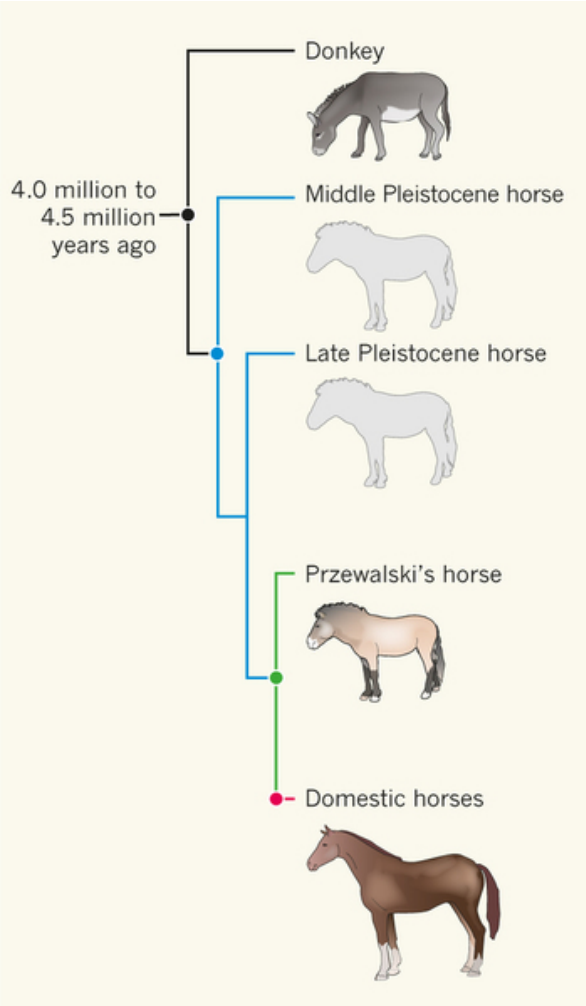
Orlando, L. et al. [2013] *Nature* **doi.org/10.1038/nature12323**

What other components of life can be ***fossilized*** and ***recovered*** with biochemical reproducibility this way?



BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501m

decoupled information

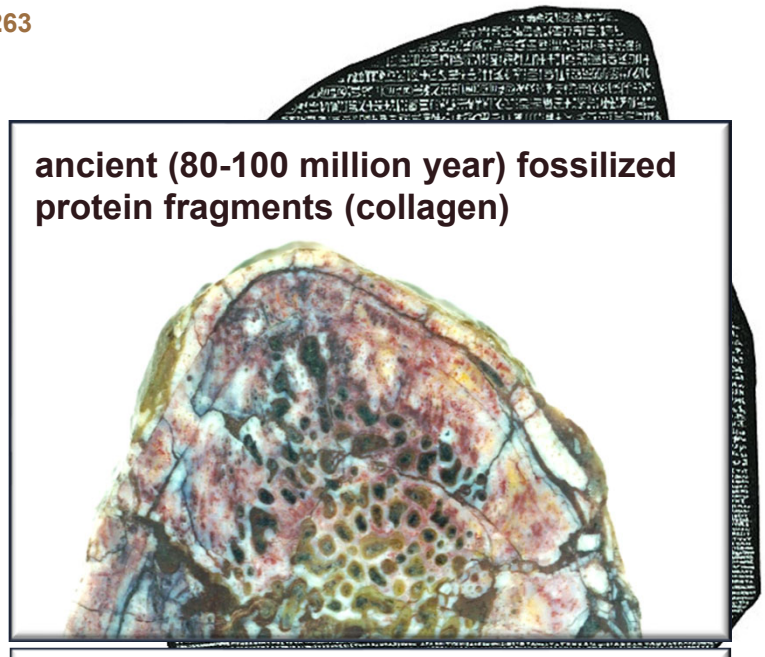


Millar & Lambert [2013]. "Ancient DNA: Towards a million-year-old genome." *Nature*. **doi:10.1038/nature12263**

Orlando, L. et al. [2013] *Nature* **doi.org/10.1038/nature12323**

What other components of life can be ***fossilized*** and ***recovered*** with biochemical reproducibility this way?

ancient (80-100 million year) fossilized protein fragments (collagen)



Schweitzer et al [2007] *Science*. **316** (5822): 277-280
Schweitzer et al [2009] *Science*. **324** (5927): 626-631.
Schroeter et al [2017] *J. Proteome Res.***16** (2):920–932
Lee et al [2017] *Nature Communications* **8**: 1422.
Service, R. [2017] *Science*. DOI: 10.1126/science.aal0679

BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501m

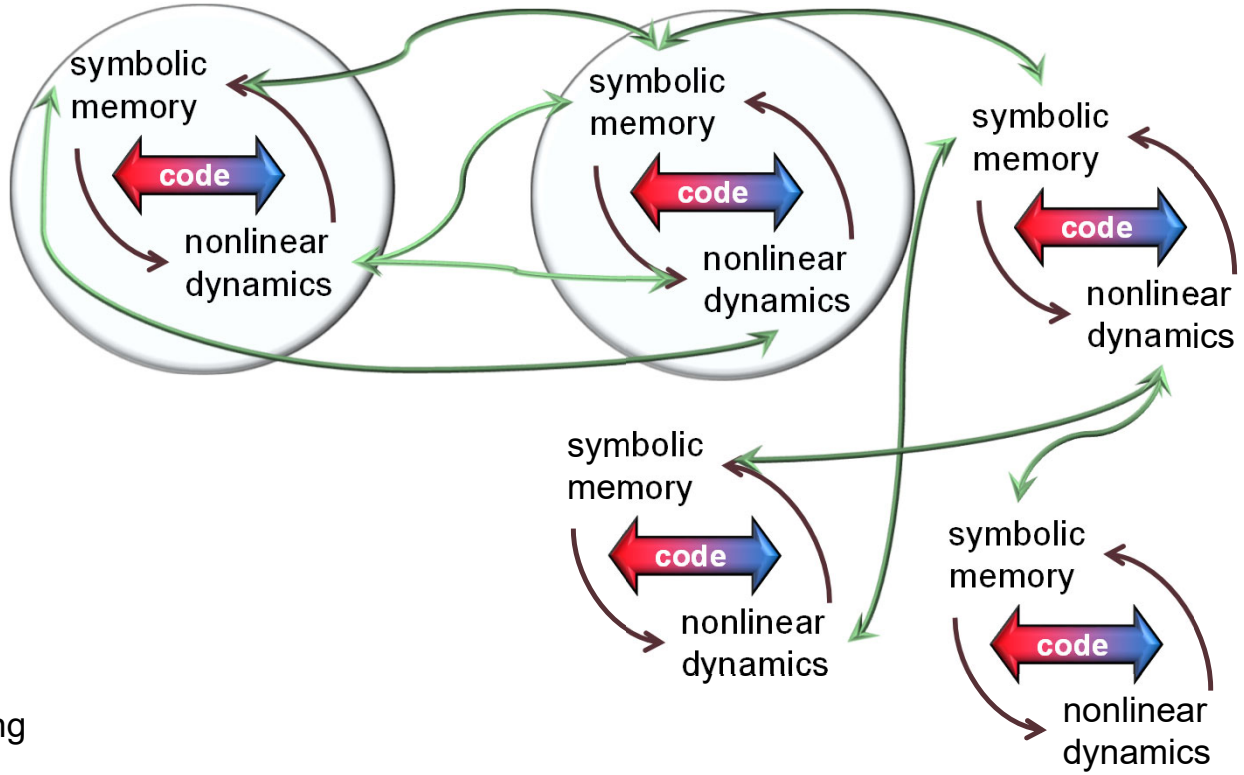
the tape is not necessarily self-contained in cells, brains, or machines

decoupling and externalization enable collective behavior

two roles of information
data/program (Turing)
passive/active (Von Neumann)
description/construction-function (Pattee)
genotype/phenotype (Biology)

"Let the whole outside world consist of a long paper tape". —John von Neumann, 1948

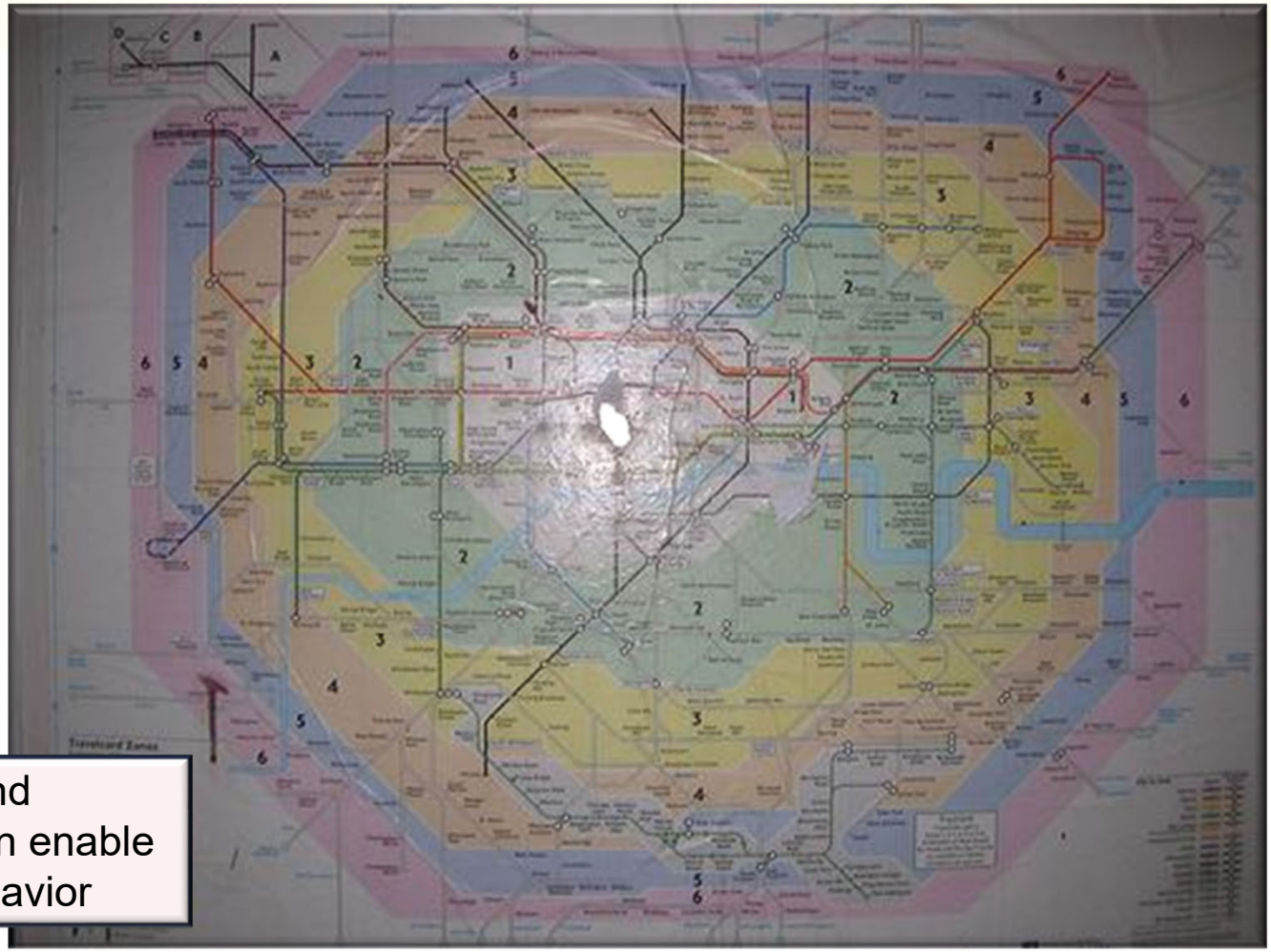
semiotic closure → **semiotic control networks**



Rocha, L.M. [2000] *Annals N.Y. Acad. Sci.* **901**(1): 207-223.
Rocha, L.M. & W. Hordijk [2005] *Artificial Life* **11**:189 - 214.

BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501m

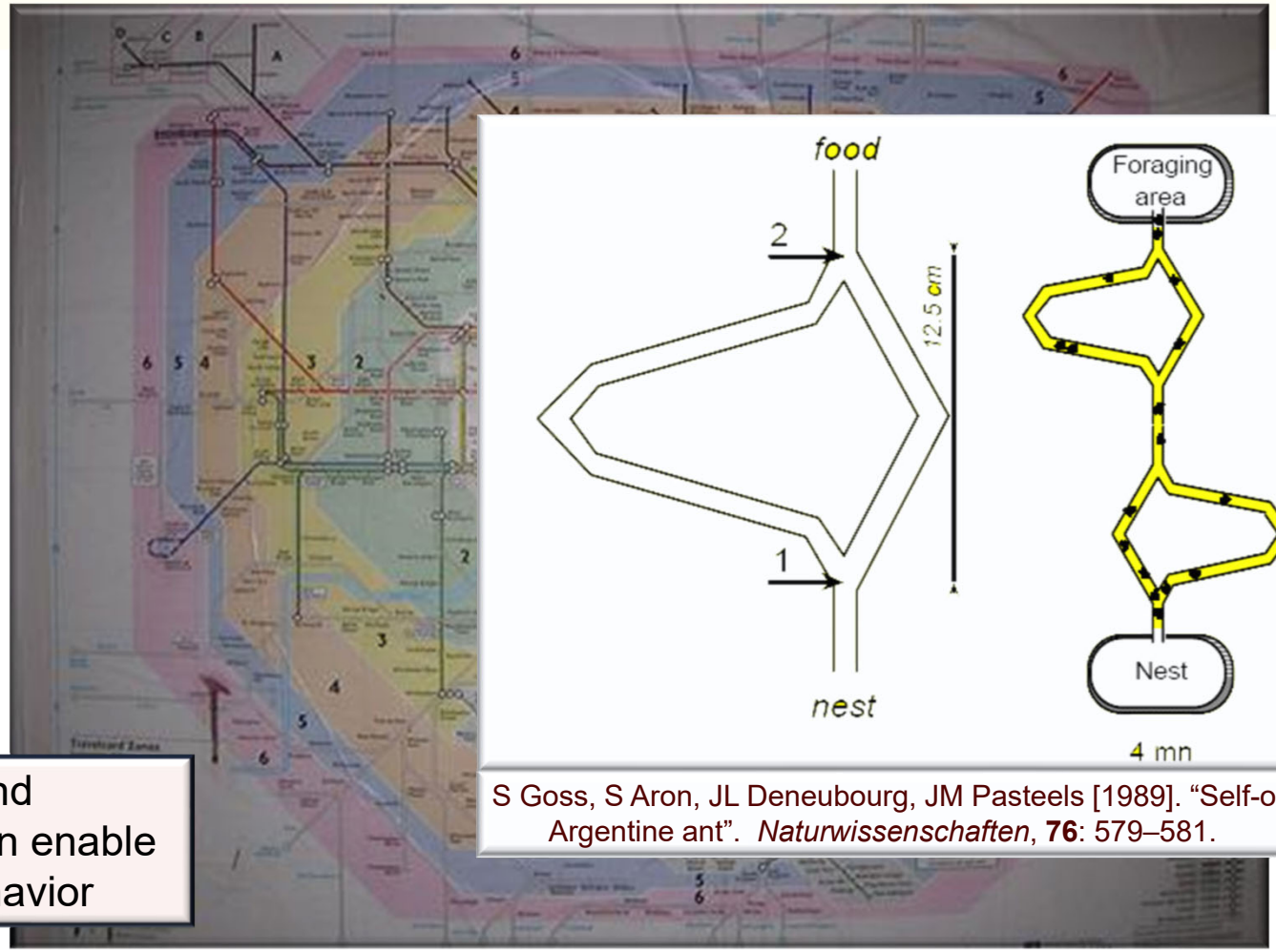
stigmergy



decoupling and externalization enable collective behavior

BINGHAMTON UNIVERSITY
STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501m

stigmergy

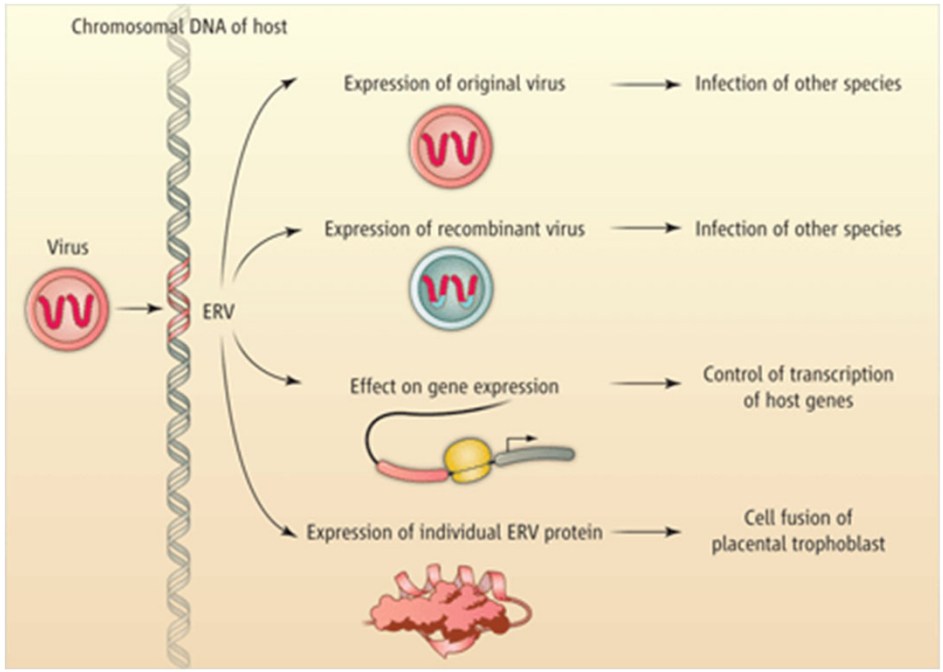


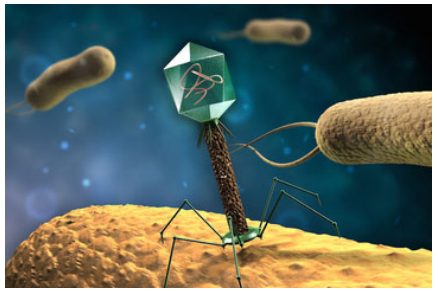
S Goss, S Aron, JL Deneubourg, JM Pasteels [1989]. "Self-organized shortcuts in the Argentine ant". *Naturwissenschaften*, **76**: 579–581.

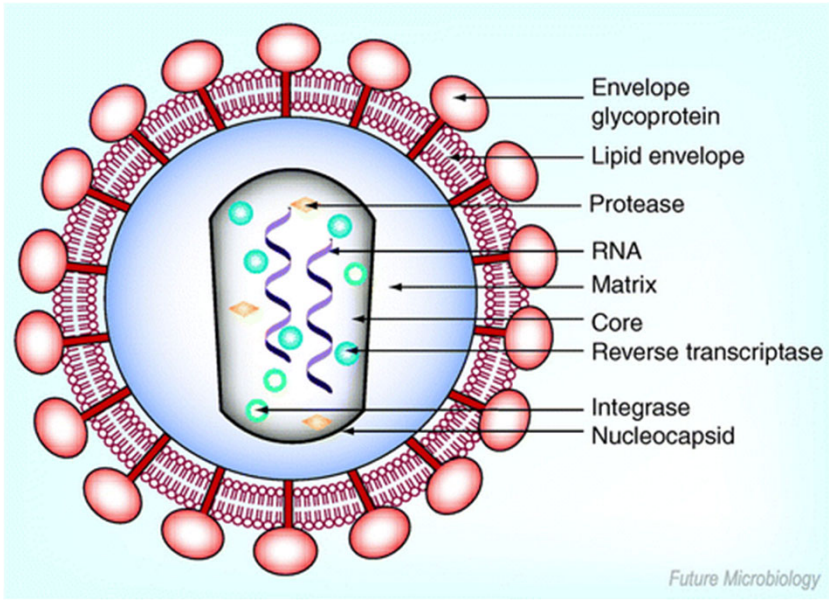
decoupling and externalization enable collective behavior

BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501m

Turing machines written on other Turing machines (naturally)







Sequences from RNA and DNA viruses found in host genomes Retroviral genomes, account **for 6 to 14% of host genomes**

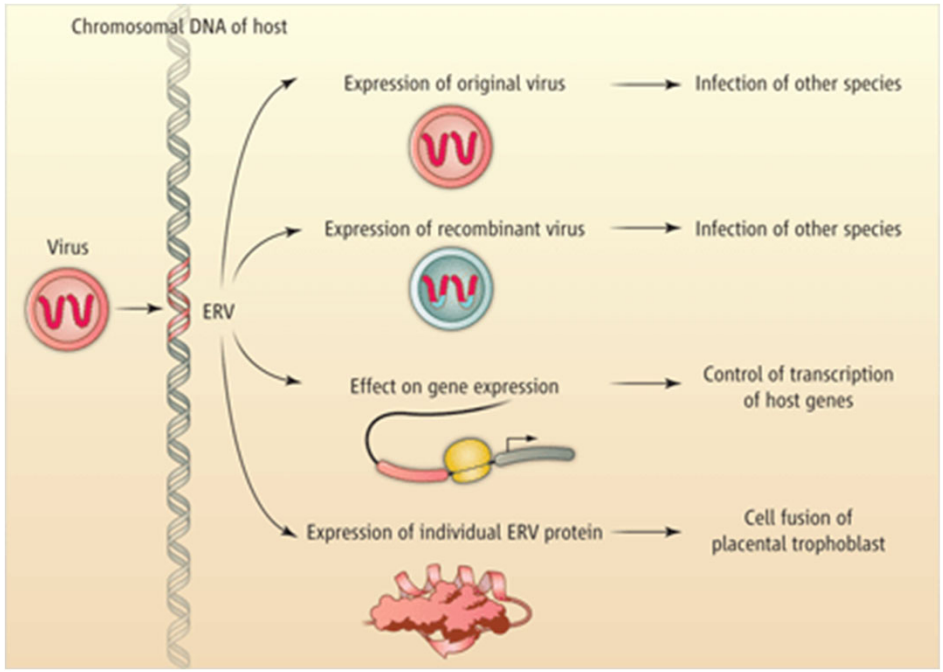
~8% of human DNA.

endogenous retroviruses (ERVs) comprise more DNA than host proteome.

Weiss & Stoye [2013]. "Our Viral Inheritance." *Science*.**340** (6134): 820-821.

BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501m

Turing machines written on other Turing machines (naturally)



Sequences from RNA and DNA viruses found in host genomes Retroviral genomes, account **for 6 to 14% of host genomes**

~8% of human DNA.

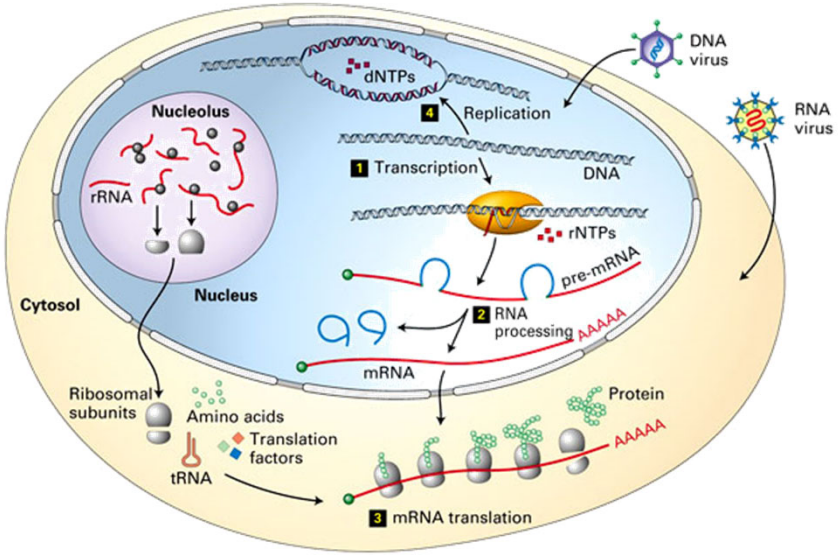
endogenous retroviruses (ERVs) comprise more DNA than host proteome.

Weiss & Stoye [2013]. "Our Viral Inheritance." *Science*.**340** (6134): 820-821.

BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501m

The social symbiome

semiotic control networks enable new, interacting levels of organization and selection, which take control of genes, organisms, and even societies.



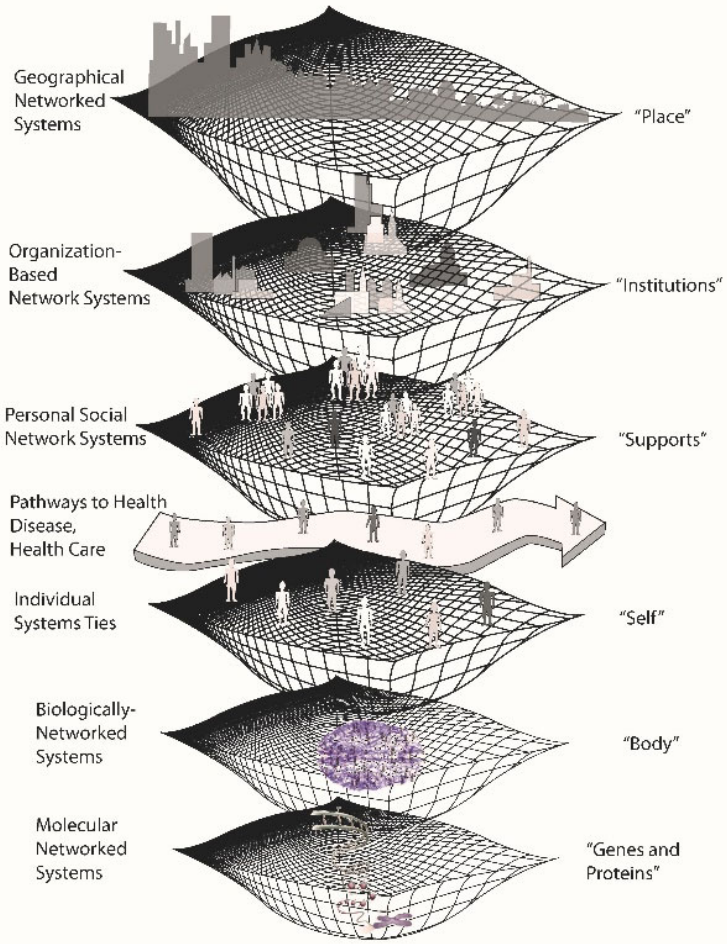
Examples: eukaryotic RNA/DNA complexity, vertebrate immunity, eusociality, cultural constraints on reproduction, GMOs (including via CRISPR), viral pandemics, etc.

BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501m

The social symbiome

semiotic control networks enable new, interacting levels of organization and selection, which take control of genes, organisms, and even societies.

Pescosolido et al [2017] ***The Social Symbiome*** Framework: Linking genes-to-global cultures in public health using network science, in The Handbook of Applied Systems Science.

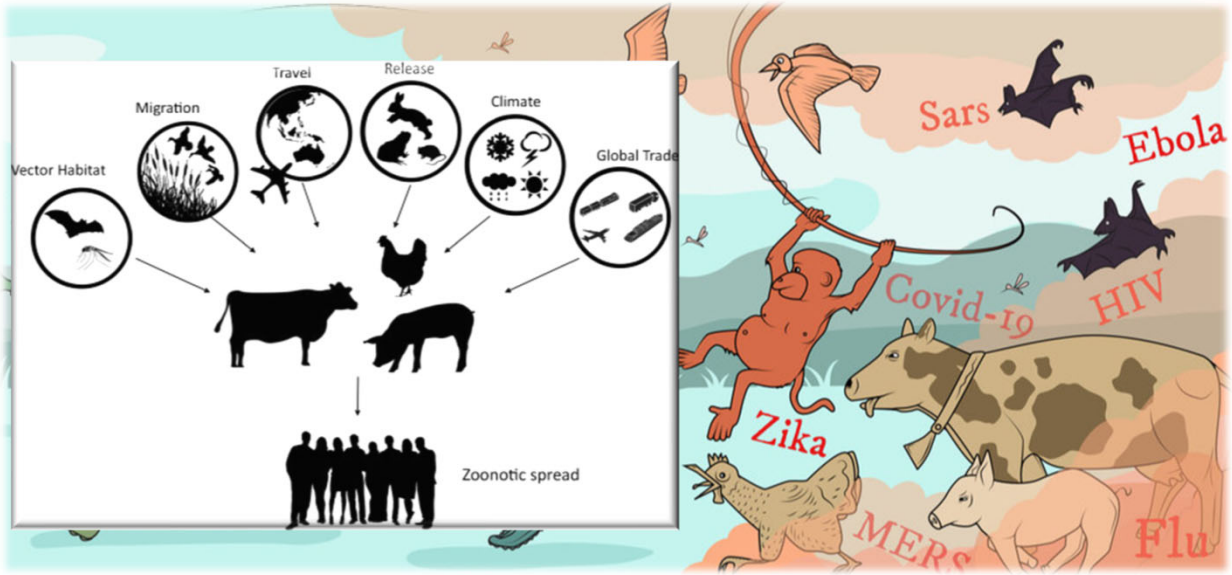


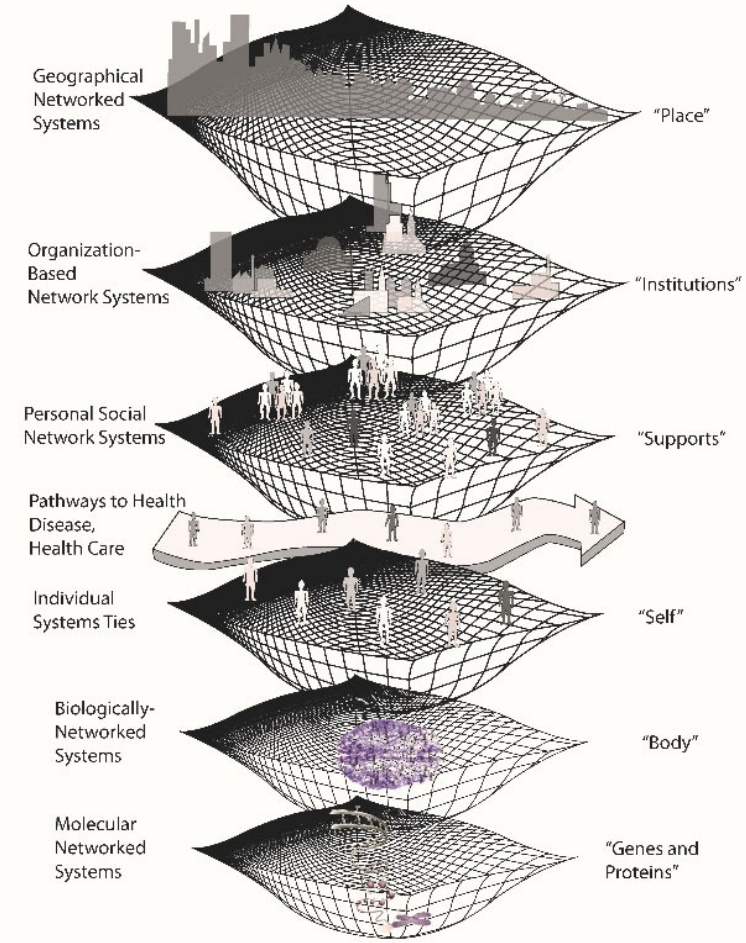
Examples: eukaryotic RNA/DNA complexity, vertebrate immunity, eusociality, cultural constraints on reproduction, GMOs (including via CRISPR), viral pandemics, etc.

BINGHAMTON UNIVERSITY
STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501m

The social symbiome

semiotic control networks enable new, interacting levels of organization and selection, which take control of genes, organisms, and even societies.



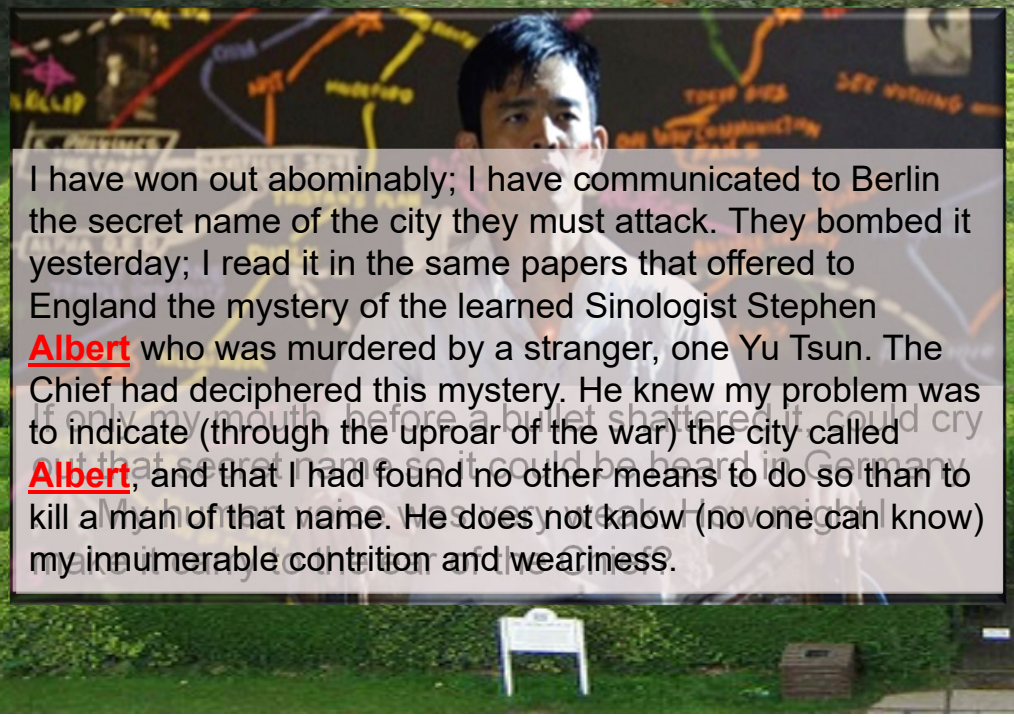


Examples: eukaryotic RNA/DNA complexity, vertebrate immunity, eusociality, cultural constraints on reproduction, GMOs (including via CRISPR), viral pandemics, etc.

BINGHAMTON UNIVERSITY
STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501m

collective behavior via the external tape

- In the presence of semiotic closure
  - Details of cells, brains, and culture can easily become irrelevant
  - Borges' garden of forking paths
    - All the knowledge on the labyrinth, but Albert is just a symbol for **control**
    - things that have (local) meaning, can easily become **irrelevant** in the intertwined semiotic control networks of collective behavior (the ***garden of forking paths***)



I have won out abominably; I have communicated to Berlin the secret name of the city they must attack. They bombed it yesterday; I read it in the same papers that offered to England the mystery of the learned Sinologist Stephen **Albert** who was murdered by a stranger, one Yu Tsun. The Chief had deciphered this mystery. He knew my problem was to indicate (through the uproar of the war) the city called **Albert**, and that I had found no other means to do so than to kill a man of that name. He does not know (no one can know) my innumerable contrition and weariness.